



May 24, 2024

The Honorable Mark R. Warner
United States Senate
703 Hart Senate Office Building
Washington, D.C. 20510

Dear Senator Warner,

Thank you for your letter dated May 14, 2024. We appreciate the opportunity to provide more details about TikTok's efforts to safeguard our community from the misuse of AI-generated content ("AIGC") and to protect election integrity.

With more than 2 billion people in over 50 countries expected to go to the polls this year, we are deeply invested in protecting the integrity of elections on TikTok. Over the last 4 years, we've worked to protect our platform through over 150 elections around the world. For more information about our commitment to election integrity, please see our [transparency center election integrity page](#) and [a recent update](#) on our ongoing work to build a safe and secure platform that remains free from outside manipulation and influence. For more information about efforts in specific countries or regions, please see our newsroom posts about TikTok's efforts in, e.g., the [United States](#), [Europe](#), [South Africa](#), and [Mexico](#).

When it comes to AI-generated content, our focus is on removing harmful content, advancing transparency so people have context about what they're viewing, and empowering responsible content creation. We enforce robust policies against misleading AI-generated content, and continue to evolve our proactive detection models while also consulting with experts and partnering with peers on shared solutions. For example, we were founding signatories to the Tech Accord to Combat Deceptive Use of AI in 2024 Elections.

To advance transparency, we require creators to label AIGC that contains realistic scenes, and built a first-of-its-kind tool to empower them do that, which has been used by more than 37 million creators globally so far. We're also proud to be the first video sharing platform to implement the [Coalition for Content Provenance and Authenticity](#) ("C2PA")'s Content Credentials technology. This means we can read Content Credentials in order to recognize and label AIGC on images and videos, and over the coming months, we'll also start attaching Content Credentials to TikTok content to help anyone using C2PA's [Verify tool](#) to identify AIGC that was made on TikTok. We are pleased to share details about these efforts and more in our responses to your questions below.

TikTok Inc.
5800 Bristol Pkwy, Suite 100
Culver City, CA 90230



- 1. What steps is your company taking to attach content credentials, and other relevant provenance signals, to any media created using your products? To the extent that your product is incorporated in a downstream product offered by a third-party, do license terms or other terms of use stipulate the adoption of such measures? To the extent you distribute content generated by others, does your company attach labels when you assess – based on either internal classifiers or credible third-party reports – to be machine generated or machine-manipulated?**

TikTok understands that while AI enables incredible creative opportunities, it can confuse or mislead viewers if they don't know content was AI-generated. Labeling helps make that context clear—which is why we [label](#) AIGC made with TikTok AI effects, and have required creators to label realistic AIGC for over a year. Our [Community Guidelines](#) include policies around edited media and AIGC, and they apply to everyone and everything on our platform, regardless of whether the content was initially generated on TikTok. When we become aware of content on our platform that violates our policies, we remove it.

Last year TikTok built a first-of-its-kind tool to help creators easily inform their community when they post AIGC, which over 37 million creators have used since last fall. This helps creators showcase the innovations behind their content and be transparent with viewers, and they can apply it to any content that has been completely generated or significantly edited by AI. This month, TikTok is expanding auto-labeling to label AIGC when it's uploaded from certain other platforms. To do this, we've joined the [Coalition for Content Provenance and Authenticity \("C2PA"\)](#) and become the first video sharing platform to implement their Content Credentials technology. Content Credentials attach metadata to content, which we can use to confidently recognize and instantly label AIGC. We started rolling out this capability on images and videos, and it will be coming to audio-only content soon.

As the first video sharing platform to implement Content Credentials, the increase in auto-labeled AIGC on TikTok may be gradual at first. As other platforms also implement Content Credentials, we'll be able to label more content. To help drive Content Credentials adoption within our industry, we have joined the Adobe-led [Content Authenticity Initiative \("CAI"\)](#).

As AI evolves, TikTok will continue advancing detection for content on our platform through these technologies. We conduct continuous, proactive searches to catch reposts of AIGC we previously removed. We will also refer to fact-checker, civil society, and other expert alerts to stay up-to-date about AIGC spreading elsewhere online.

- 2. What specific public engagement and education initiatives have you initiated in countries holding elections this year? What has the engagement rate been thus far and what proactive steps are you undertaking to raise user awareness on the availability of new tools hosted by your platform?**

As mentioned in our response to Question 1, labeling can be a helpful tool for clarifying context around AIGC. While experts widely recommend AIGC labeling as a way to



support responsible content creation, they also caution that labels can cause confusion if viewers don't have context about what they mean. That's why we've been working with experts to develop media literacy campaigns that can help our community identify and think critically about AIGC and misinformation. Working with [MediaWise](#), a program of the Poynter Institute, we'll release 12 videos throughout the year that highlight universal media literacy skills while explaining how TikTok tools like AIGC labels can further contextualize content. We'll also be launching a campaign to raise awareness around AI labeling and potentially misleading AIGC, with a series of videos we developed with expert guidance from [WITNESS](#).

Additionally, TikTok recently launched an updated [Harmful Misinformation Guide](#) within our [Safety Center](#) to share information about the rules, tools and tactics that can help our community have a trustworthy experience on TikTok and beyond. We also regularly refresh materials within our [Transparency Center](#) to share the latest information on our efforts to [support responsible and transparent AI-generated content](#), [combat harmful misinformation](#), and [protect the integrity of elections](#), for instance. TikTok is also localizing educational content in certain markets ahead of elections, for example, by launching media literacy campaigns in 18 European countries (with nine additional campaigns scheduled to go live in 2024), and by providing [video](#) and [other resources](#) as part of an AIGC awareness campaign in Mexico. Additional information on our elections-related media literacy initiatives can be found within newsroom posts on our efforts to prepare for the 2024 elections in, e.g., the [United States](#), [Europe](#), [South Africa](#), and [Mexico](#).

3. What specific resources has your company provided for independent media and civil society organizations to assist in their efforts to verify media, generate authenticated media, and educate the public?

In addition to the educational efforts referenced in our response to Question 2 and our implementation of C2PA Content Credentials as referenced in our response to Question 1, TikTok supports [C2PA](#)'s working groups as a General Member. TikTok has also joined the [CAI](#) to encourage wider adoption of Content Credentials and responsible use of generative AI tools. CAI produced a Media Literacy [curriculum](#) with age-appropriate lesson plans for middle school, high school, and higher education students to develop critical media and visual literacy skills.

TikTok partners with experts across the world to support consistent and accurate moderation, understand local context, and empower our community with authoritative information. Through TikTok's Global Fact-Checking program, we work closely with 18 [IFCN-accredited](#) fact-checking organizations to assess the accuracy of content and support informed, responsible moderation decisions. Fact-checkers do not moderate, action, or label content directly on TikTok. Instead, they provide factual assessments that empower our moderators and policy teams to apply our misinformation policies accurately. Together, our 18 global fact-checking partners cover more than 50 languages and 100 markets globally.

We've also launched [Research Tools](#) that provide eligible researchers in the U.S. and Europe with access to public data on content and accounts on our platform in a way that



preserves individual user privacy. We're working to provide increased access to the Research Tools in the future.

Additionally, our global Community Partner Channel provides selected organizations an additional route for reporting content that they believe violates our Community Guidelines so that it can be reviewed by our teams. To date, more than 250 organizations who specialize in a range of safety issues use our Community Partner Channel. As explained in our response to Question 7, anyone can [report](#) content in-app for review.

Media literacy is also integral to our elections strategy, and we collaborated with local creators and civil society to promote media literacy best practices that reached millions more people on TikTok last year. Throughout 2024, we'll continue to partner with experts and fact-checking organizations around the world to deliver engaging media literacy campaigns about misinformation, identifying AIGC, and more.

4. What has been your company's engagement with candidates and election officials with respect to anticipating misuse of your products, as well as the effective utilization of content credentialing or other media authentication tools for their public communications?

For years, TikTok has not allowed [paid political advertising](#), and accounts belonging to [politicians or political parties](#) are not eligible to advertise or make money on TikTok. Our political advertising policy includes both traditional paid advertisements and creators receiving compensation to support or oppose a candidate for office. Furthermore, TikTok does not allow content that shares or shows fake authoritative sources or crisis events, or falsely shows public figures such as politicians in certain contexts, as even when appropriately labeled, AIGC or edited media in these contexts may still be harmful. This includes being bullied, making an endorsement of a political view, or being endorsed.

We also recognize that accounts belonging to news organizations, politicians, political parties and governments play a unique role in civic discourse, and apply more nuanced account enforcement policies to protect the public interest, which we explain in our [Community Guidelines](#).

TikTok partners with electoral commissions and fact-checking organizations to build Election Centers that connect people to trustworthy information about voting. We've launched our U.S. Elections Center, in partnership with nonprofit [Democracy Works](#), which provides our 170 million U.S. community members with reliable voting information for all 50 states and Washington, D.C. We direct people to the Elections Center through prompts on relevant election content and searches.

In addition, we're continuing to build features that provide additional context about content and accounts on TikTok. For example, blue "verified" checks confirm that notable accounts are who they say, and in the U.S., we require [government, politician, and political party accounts](#) to be verified. We label content that our fact-checkers



determine as unsubstantiated and will expand media literacy resources to these labels this year.

5. Has your company worked to develop widely-available detection tools and methods to identify, catalogue, and/or continuously track the distribution of machine-generated or machine-manipulated content?

Over the coming months, we'll start attaching C2PA's Content Credentials to TikTok content, which will remain embedded in content when downloaded. That means that anyone will be able to use C2PA's Verify tool to help [identify](#) AIGC that was made on TikTok and even learn when, where and how the content was made or edited. Other platforms that adopt C2PA's Content Credentials will be able to read that metadata so they can automatically label it.

6. (To the extent your company offers social media or other content distribution platforms) What kinds of internal classifiers and detection measures are you developing to identify machine-generated or machine-manipulated content? To what extent to these measures depend on collaboration or contributions from generative AI vendors?

TikTok continues to develop and test models to assist in the detection and moderation of AIGC uploaded to the platform. Additionally, along with implementing C2PA's Content Credentials technology, we are working with other members of C2PA and CAI to expand options for reading and writing metadata.

7. (To the extent your company offers social media or other content distribution platforms) What mechanisms has your platform implemented to enable victims of impersonation campaigns to report content that may violate your Terms of Service? Do you maintain separate reporting tools for public figures?

TikTok recognizes that AIGC or edited media may still be harmful even when appropriately labeled. Our [Community Guidelines](#) do not allow content that falsely shows public figures in certain contexts. This includes being bullied, making an endorsement, or being endorsed. We are also committed to protecting people's privacy, and we do not allow content that contains the likeness of young people, or the likeness of adult private figures used without their permission.

We provide options for our community to report content that they believe violates our Community Guidelines and [Advertising Guidelines](#). Content that contains deepfakes or synthetic media that are potentially in violation of our policies can be reported within the TikTok app by:



1. Going to the video you need to report.
2. Pressing and long holding on the video.
3. Selecting Report, tapping on 'Misinformation' and then 'Deepfakes, synthetic media, and manipulated media.'

TikTok's [Community Guidelines](#) also do not allow account behavior that may spam or mislead our community, and this includes impersonation accounts. Users can report account impersonations in-app and through the [TikTok website](#).

If our community is watching content from an account belonging to a government, politician, or political party, we want them to know the account is genuine. [Verification](#) lets our community know an account is authentic and belongs to the user it represents, which is a way to build trust between high-profile creators and their community. While many political accounts have added the verified badge to their profile already, in the United States we have a mandatory verification [policy](#) for accounts belonging to governments, politicians, and political parties.

- 8. (To the extent your company offers generative AI products) What mechanisms has your platform implemented to enable victims of impersonation campaigns that may have relied on your models to report activity that may violate your Terms of Service?**

Please see our response to Question 7 for information about reporting impersonation activity.

- 9. (To the extent your company offers social media or other content distribution platforms) What is the current status of information sharing between platforms on detecting machine-generated or machine-manipulated content that may be used for malicious ends (such as election disinformation, non-consensual intimate imagery, online harassment, etc.)? Will your company commit to participation in a common database of violative content?**

We know that no platform can solve the challenges around AIGC alone, so TikTok works with others on interoperable solutions. For example, TikTok supported the launch of the Partnership on AI ("PAI")'s [Framework for Responsible Practices for Synthetic Media](#). A [case study](#) on PAI's website details how TikTok launched new AI labeling policies to prevent misleading content and empower responsible creation.

In addition to our efforts combat the deceptive use of AI with respect to elections, TikTok also participates in initiatives to prevent the misuse of AIGC in other contexts. For example, we partnered with StopNCII.org to bolster efforts to stop the spread of non-consensual intimate imagery. This initiative allows people to report non-consensual sexual images (also referred to as image-based sexual abuse). StopNCII translates those images to hashes and shares them with TikTok and other companies to remove them from the app.



TikTok welcomes the opportunity to join additional initiatives and to continue working with other companies on these important efforts.

We thank you for your questions and appreciate the opportunity to provide additional details and context around our efforts to safeguard our community from the misuse of AIGC and to protect election integrity on our platform. We welcome a continued dialogue with your office regarding our shared priority of maintaining a safe platform for our users.

Sincerely,

A handwritten signature in black ink, appearing to read "Michael Beckerman".

Michael Beckerman
Vice President and Head of Public Policy, Americas, TikTok